

人口集聚的消费效应

蔡郑宇 颜 宇

目录

附录 I 人口密度与交易成本.....	1
附录 II 数据来源	3
附录 III 变量说明	4
附录 IV 表 2 详细结果	7
附录 V 稳健性检验	11
附录 VI 异质性分析	13
附录 VII 人口集聚与不可贸易服务密度	15
附录 VIII 住房支出估算	16
参考文献	17

附录I 人口密度与交易成本

消费匹配的质量和机会受到人口密度的影响。消费匹配的质量是指消费者对商品特征的特异性需求与市场中提供的商品特征间存在的差距。人口密度越高，消费匹配的质量越高，消费者的商品特征需求与市场中的商品特征差异越小。我们可以用城市劳动力市场中常用的单位圆模型(Duranton and Puga, 2004; Helsley and Strange, 1990)来进行解释。假设有 n 个消费者，每个消费者对应一种独有的商品特征需求。这些异质性的商品特征在一个单位圆上的坐标为 y 并服从均匀分布， n 即为此单位圆空间中的人口密度。在同一空间中，假设有 m 个异质性的厂商。每个厂商根据消费者的需求提供一个具有独有特征的商品，这些商品特征在单位圆上的坐标为 x 。消费匹配的质量即厂商提供的商品特征与消费者需要的商品特征间的距离， $|x - y|$ ，距离越小匹配质量越高。不同的消费者可以消费同一种特征的商品，令 $Y(x)$ 为消费商品特征 x 的所有消费者的坐标集。根据Helsley and Strange (1990)的证明可知，对于商品特征 x 总的特征距离期望为 $E \sum_{Y(x)} |x - y| = n/4m^2$ 。易知每家厂商提供的商品特征的市场范围 θ 是一致的，则 $E \sum_{Y(x)} |x - y| = \theta^2/4n$ 。显然此式对于 n 的一阶导数为负，因此人口密度越大，特征距离的总期望越小，消费匹配质量的期望越高。从消费匹配质量的角度来看， $\tau(d_j)$ 是 d_j 的减函数。

我们提供了一个商品特征匹配的数值案例。假设每一个消费者对应一个独有的商品特征偏好，这些特征偏好均匀地分布在一个单位圆上。下图中的图 I 1 展示了一个有四个消费者的情况。这四位消费者或者说四种异质性商品特征按顺时针均匀地分布在一个单位圆上，为分析简便，假设消费者 1 的坐标为 0。消费者的坐标集为 $\{0, 1/4, 1/2, 3/4\}$ 。厂商进入市场并提供独有特征的商品，假设每个厂商的市场范围可以覆盖两个消费者，那么此例中厂商数量为 2。厂商 1 提供的商品特征坐标为 $1/8$ ，厂商 2 提供的商品特征坐标为 $5/8$ 。每个消费者面对的特征距离均为 $1/8$ ，每一种商品对应的总特征距离为 $1/4$ 。假设消费者数量增加到 6，其它条件不变。由图 I 2 可知，消费者的坐标集为 $\{0, 1/6, 1/3, 1/2, 2/3, 5/6\}$ ，有三家厂商分别提供特征为 $\{1/12, 5/12, 3/4\}$ 的商品。每个消费者面对的特征距离变为 $1/12$ ，每一种商品对应的总特征距离为 $1/6$ 。图 I 3 展示了消费者数量增加到 12 的情况，厂商增加到 6 家。每个消费者面对的特征距离为 $1/24$ ，每一种商品对应的总特征距离为 $1/12$ 。由此可见，随着消费者密度的增加，特征距离越小，消费匹配的质量越高，消费者对更符合偏好的商品特征支付意愿更高，因此消费支出提高。

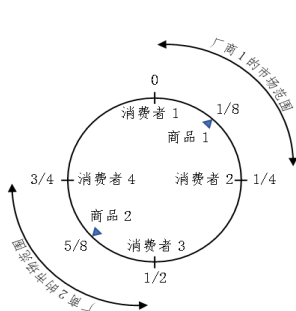


图 I 1 消费者密度为 4

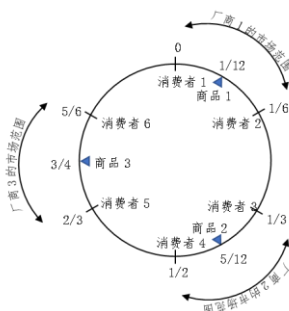


图 I 2 消费者密度为 6

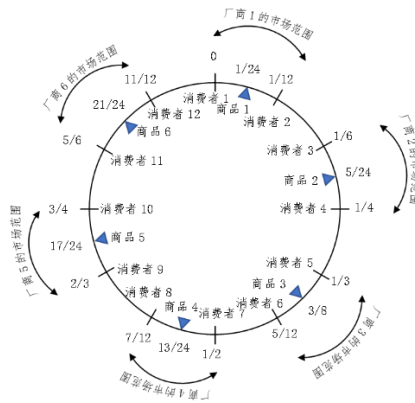


图 13 消费者密度为 12

另一方面，人口密度可能会影响消费匹配的机会。如果将交易的达成视为买方与卖方的成功匹配，那么一个地区的买方与卖方的数量越多，消费交易达成的机会越大。Duranton and Puga (2004) 对解释这一机制的搜寻匹配模型进行了综述。虽然这类模型都是建立在劳动者求职的情景上的，但 Duranton and Puga (2004) 也提到这类模型本质上是对买方与卖方的匹配进行建模。第一支文献研究非协调随机匹配。这种方法得到的总体匹配函数是规模报酬递减的，但是随着交易双方的数量增加规模报酬递减的程度会降低，在极限下趋近规模报酬不变。第二支文献关注当市场参与者选择他们的位置以更好地对搜寻目标进行定位时，如何内生地产生搜索摩擦，此模型得到的总体匹配函数也是规模报酬不变的。在第三类模型中，Coles (1994)、Coles and Smith (1998) 得到了规模报酬递增的总体匹配函数，这表明消费行为会集聚在市场参与者密度大的地方。因此，虽然人口密度会提高消费匹配机会，但有可能是规模报酬不变的。即便如此，由于消费匹配质量的提高，人口密度增加也会增加家庭消费的总支出。

人口集聚产生的负外部性可能会增加交易成本，抑制消费。人口密度增加引发交通拥堵，造成通勤时间的增加。通勤时间增加可能会挤出闲暇时间，间接减少个体的消费时间和消费需求。有文献表明，人口密集的城市有较高的犯罪率 (Glaeser and Sacerdote, 1999)。犯罪率的增加会扭曲家庭消费决策，使得家庭可见性商品的消费显著下降 (Mejía and Restrepo, 2016)。人口密度较高的城市可能导致更为严重的总体污染 (Gaigné et al., 2012)。污染的加剧导致人们不愿外出购买商品和服务，从而影响消费行为。有证据发现，空气污染使得线下消费显著下降，同时这一负面影响不太可能被网上购物所抵消 (Kang et al., 2019)。综上所述， $\tau(d_j)$ 可能是非线性的，但与集聚经济的非线性类似 (Glaeser and Gottlieb, 2008)，其具体的函数形式是难以确定的。

附录II 数据来源

本文使用的主要数据来源于西南财经大学中国家庭金融调查与研究中心 2013 至 2017 年在全国范围内开展的三轮中国家庭金融调查 (CHFS)。CHFS 数据获取了居民家庭的人口统计特征、家庭财富、收入支出、保险保障及就业方面的信息,详细记录了家庭的各项收入和消费状况,为研究居民消费问题提供了可靠的数据支持 (尹志超等, 2015)。

被解释变量的数据来源是 CHFS 的家庭消费和收入数据。由于 CHFS 收集的是受访家庭调查年份上一年度的消费和收入数据,即 2015 和 2017 年的数据是 2014 和 2016 年产生的。为处理方便,我们使用 2015 和 2017 年的数据 (实际为 2014 和 2016 年发生) 来构造被解释变量,使用同一追踪家庭 2013 和 2015 年的数据来构造控制变量。主要解释变量城市人口密度的计算也是使用的 2013 和 2015 年的数据。因此,本文回归模型中的控制变量均采用滞后一期的形式引入模型。

城市人口数据来自 LandScan 人口数据库。LandScan 全球人口分布数据库利用高质量人口普查数据,基于空间数据、图像分析和多变量的分区建模法,并结合行动轨迹和集体出行习惯估计常住人口的空间分布,已被广泛应用于人口数据的研究 (Li and Liu, 2018)。与行政统计数据相比, LandScan 人口数据是具有高分辨率的栅格数据,其优势在于方便研究者使用地理信息系统软件提取不同地理尺度下的人口数据,以满足不同的研究需求。由于城市与非城市地区可能存在系统性的消费行为差异,本文所研究的“城市”指的是地级及以上城市的市辖区,而市辖区的常住人口行政统计数据难以获取。于是我们利用 LandScan 人口数据的上述特性获取了市辖区的常住人口数据。LandScan 全球人口数据可能存在的局限在于其模型参数的校准并不完全符合我国的基本事实,因此其估计出的人口数据可能会存在非随机偏误,而这种测量误差会导致本研究的估计值出现趋零偏误。为了解这种潜在偏误对实证结果的影响,我们在稳健性检验中也使用了行政统计数据计算的地级市人口密度。地级市常住人口数据来自《中国区域经济统计年鉴》与 CEIC 中国经济数据库。城市面积数据来自《中国城市统计年鉴》。

此外,数字普惠金融的数据来源于北京大学数字金融研究中心。机制检验中的相关数据来源于高德地图 POI 数据,中国综合社会调查 (CGSS),和聚汇数据。居民消费价格指数的数据来源于各省统计年鉴、各城市统计年鉴、各城市的国民经济和社会发展统计公报。具体变量对应的数据来源见附录 III。

附录III 变量说明

主要被解释变量中,家庭人均消费支出由家庭总消费除以家庭成员数然后取自然对数来衡量;家庭平均消费倾向由家庭总消费除以家庭可支配收入得到。所有的名义变量都使用以2010年为基期的省级居民消费价格指数换算为实际值。家庭总消费包括食品支出、衣着支出、居住支出(包括物业管理费、暖气费和住房装修维修支出)、生活用品及服务支出、耐用品支出、教育娱乐支出、交通通信支出、医疗保健支出和其他支出。

关键解释变量城市人口密度由城市市辖区的人口规模与城市市辖区的建成区面积之比来度量。LandScan数据集显示了每个栅格上的常住人口数量,栅格分辨率为30"。城市市辖区内栅格的人口总量就是城市人口规模。建成区是城市实际征用的用于非农业生产和生活的土地,是城市人口、经济活动所在地,因此我们参考钟粤俊等(2020)使用建成区的面积去计算城市人口密度。

模型中的其它控制变量有收入资产特征变量,包括:家庭人均收入、家庭人均金融资产、是否拥有住房、家庭人均住房资产、是否有住房负债;家庭特征变量,包括:是否自营工商业、人口结构(包括少年抚养比、老年抚养比、无工作家庭成员占比)、受访者及配偶中拥有养老保险人数、受访者及配偶中拥有医疗保险人数;户主特征变量,包括:风险偏好、性别、年龄、受教育年限、婚姻状况、健康状况、户籍类型。同时,本文还尝试控制了可能影响家庭消费的城市层面变量,包括国内生产总值(GDP)规模、第三产业占GDP比重、金融发展水平、数字普惠金融、住房价格。

具体的变量说明如表III1所示。

表III1 变量说明

变量名	变量定义	数据来源
被解释变量		
家庭人均消费(元)	家庭总消费(包括食品支出、衣着支出、居住支出(包括物业管理费、暖气费和住房装修维修支出)、生活用品及服务支出、耐用品支出、教育娱乐支出、交通通信支出、医疗保健支出和其他支出,下同)除以家庭规模	CHFS
家庭平均消费倾向	家庭总消费除以家庭可支配收入	CHFS
主要解释变量		
人口密度(人/平方公里)	城市市辖区人口规模(城市市辖区内栅格的人口总量)除以城市市辖区建成区面积	LandScan数据集、城市统计年鉴
家庭资产变量		
家庭人均收入(元)	家庭可支配收入除以家庭规模	CHFS
家庭人均金融资产(元)	家庭总金融资产(包括活期存款、定期存款、股票、基金、理财产品、债券、衍生品、非人民币资产、黄金、其他金融资产、现金、借出款)除以家庭规模	CHFS
是否拥有住房	0-1变量,0为无住房,1为有住房	CHFS
家庭人均住房价值(元)	家庭总住房价值除以家庭规模	CHFS
是否有住房负债	0-1变量,0为无住房负债,1为有住房负债	CHFS
家庭特征变量		
是否自营工商业	0-1变量,0为不从事工商业,1为从事工商业	CHFS
少年抚养比	小于及等于14周岁的家庭成员数除以家庭规模	CHFS
老年抚养比	大于及等于65周岁的家庭成员数除以家庭规模	CHFS

失业成员占比	无工作家庭成员数除以家庭规模	CHFS
受访者及配偶中 养老保险人数（人）	受访者及配偶中有养老保险的人数	CHFS
受访者及配偶中 医疗保险人数（人）	受访者及配偶中有医疗保险的人数	CHFS
户主特征变量		
风险偏好	风险偏好通过如下问题测度：如果您有一笔钱，您愿意选择哪种投资项目？选项为：1. 高风险、高回报的项目；2. 略高风险、略高回报的项目；3. 平均风险、平均回报的项目；4. 略低风险、略低回报的项目；5. 不愿意承担任何风险。我们根据选项 1 到 5 依次将风险偏好变量赋值为 5、4、3、2、1，该数值越大表明越偏好风险	CHFS
户主性别	0-1 变量，0 为女性，1 为男性	CHFS
户主年龄（岁）	调查年份减去出生年份	CHFS
户主受教育年限（年）	根据受教育程度的类别进行折算，0 为没上过学，6 为小学，9 为初中，12 为高中、中专/职高，15 为大专/高职，16 为大学本科，19 为硕士研究生，22 为博士研究生	CHFS
户主婚姻状况	0-1 变量，0 为未婚，1 为已婚	CHFS
户主健康状况	范围在 0-5，1 为非常不好，2 为不好，3 为一般，4 为好，5 为非常好	CHFS
户主户籍类型	0-1 变量，0 为农业户籍，1 为非农户籍	CHFS
城市特征变量		
GDP 规模（元）	城市 GDP 规模	城市统计年鉴
第三产业占 GDP 比重	第三产业 GDP 除以城市 GDP 规模	城市统计年鉴
金融发展水平	年末金融机构各项贷款余额除以 GDP 规模	城市统计年鉴
住房价格（元/平方米）	家庭住房市场价值除以住房建筑面积，然后在城市市辖区层面使用家庭调查权重进行加权平均	CHFS
数字普惠金融	具体编制过程见郭峰等（2020）	北京大学数字金融研究中心
其它变量		
不可贸易服务的匹配质量	不可贸易服务的信息熵（信息熵的具体定义参见正文中描述，下同）除以城市市辖区建成区面积	高德地图 POI 数据、城市统计年鉴
餐饮服务的匹配质量	餐饮服务的信息熵除以城市市辖区建成区面积	高德地图 POI 数据、城市统计年鉴
购物服务的匹配质量	购物服务的信息熵除以城市市辖区建成区面积	高德地图 POI 数据、城市统计年鉴
科教文化服务的匹配质量	科教文化服务的信息熵除以城市市辖区建成区面积	高德地图 POI 数据、城市统计年鉴
生活服务的匹配质量	生活服务的信息熵除以城市市辖区建成区面积	高德地图 POI 数据、城市统计年鉴
体育休闲服务的匹配质量	体育休闲服务的信息熵除以城市市辖区建成区面积	高德地图 POI 数据、城市统计年鉴
社交频率	CGSS 会询问受访者“您与邻居进行社交娱乐活动的频繁程度”、“您与其他朋友进行社交娱乐活动的频繁程度”，根据	CGSS

	问题选项从“从不”到“几乎每天”分别赋值为 1-7，取两个问题的均值代表受访者的社交频率，数值越大表示社交活动越频繁	
总人情支出（元）	家庭在传统节假日以及红白喜事中，支付给非家庭成员礼金、礼品的实际货币价值	CHFS
人均人情支出（元）	家庭在传统节假日以及红白喜事中，支付给非家庭成员礼金、礼品的实际货币价值，除以家庭规模	CHFS
新房价格（元/平方米）	地级市住宅销售总额除以住宅销售总面积	CEIC 中国经济数据库
二手房价格（元/平方米）	以区人口占比（区人口规模除以城市市辖区人口规模）为权重，将区二手房价格在城市市辖区层面进行加权平均	LandScan 数据集、聚汇数据
工具变量 IV1	城市建成区面积与距长江、黄河或海岸线的最近距离的乘积，乘以家庭所在省份除自身城市以外的其他城市人口的年变化率	城市统计年鉴、中国科学院资源与环境科学数据中心、LandScan 数据集
工具变量 IV2	1953 年城市人口密度的线性时间趋势	1953 年第一次全国人口普查数据、城市统计年鉴

附录IV 表 2 详细结果

表IV 1 表 2 详细结果

	A. 被解释变量：家庭平均消费倾向				
	OLS		FE		FE+时间趋势
	(1)	(2)	(3)	(4)	(5)
Ln(人口密度)	0.017 (0.061)	0.040 (0.040)	0.357*** (0.122)	0.430*** (0.134)	0.404*** (0.132)
Ln(家庭人均收入)		-0.531*** (0.017)	-0.688*** (0.026)	-0.689*** (0.026)	-0.696*** (0.028)
Ln(家庭人均金融资产)		0.037*** (0.006)	0.007 (0.006)	0.008 (0.006)	0.006 (0.008)
Ln(家庭人均住房价值)		0.009*** (0.003)	0.000 (0.002)	0.001 (0.002)	-0.003 (0.003)
少年抚养比		-0.415*** (0.101)	-0.391** (0.175)	-0.403** (0.177)	-0.353* (0.186)
老年抚养比		-0.215*** (0.062)	0.128 (0.140)	0.123 (0.139)	0.155 (0.132)
失业成员占比		0.241*** (0.051)	0.064 (0.071)	0.068 (0.071)	0.046 (0.079)
Ln(城市GDP)				0.101 (0.156)	0.098 (0.150)
第三产业占比				-0.040 (0.496)	-0.143 (0.553)
金融发展水平				0.064*** (0.022)	0.069** (0.034)
Ln(住房价格)				-0.056 (0.041)	-0.034 (0.040)
Ln(数字普惠金融水平)				-0.216 (0.278)	0.187 (0.529)
Ln(家庭人均收入)的时间趋势					-0.001 (0.012)
Ln(家庭人均金融资产)的时间趋势					-0.002 (0.004)
Ln(家庭人均住房价值)的时间趋势					-0.002 (0.002)
少年抚养比的时间趋势					0.048 (0.059)
老年抚养比的时间趋势					0.061* (0.033)
失业成员占比的时间趋势					-0.023 (0.028)
是否拥有住房的时间趋势					-0.030

					(0.028)
是否有住房负债的时间趋势					-0.050*
					(0.026)
是否自营工商业的时间趋势					-0.032
					(0.027)
受访者及配偶中养老保险人数为1					0.023
的时间趋势					(0.039)
受访者及配偶中养老保险人数为2					0.011
的时间趋势					(0.029)
受访者及配偶中医疗保险人数为1					-0.086**
的时间趋势					(0.040)
受访者及配偶中医疗保险人数为2					-0.023
的时间趋势					(0.033)
Ln(城市GDP)的时间趋势					0.001
					(0.013)
第三产业占比的时间趋势					-0.017
					(0.125)
金融发展水平的趋势					0.005
					(0.015)
Ln(住房价格)的时间趋势					0.017
					(0.017)
常数项	0.892	5.344***	4.135***	3.342	-99.237
	(0.580)	(0.428)	(1.232)	(4.034)	(390.401)
样本数	10883	10883	10883	10883	10883
B. 被解释变量: Ln(家庭人均消费)					
	OLS		FE		FE+时间趋势
	(1)	(2)	(3)	(4)	(5)
Ln(人口密度)	-0.101	0.004	0.271**	0.355**	0.343**
	(0.178)	(0.065)	(0.119)	(0.137)	(0.146)
Ln(家庭人均收入)		0.184***	0.110***	0.110***	0.116***
		(0.012)	(0.011)	(0.011)	(0.016)
Ln(家庭人均金融资产)		0.060***	-0.004	-0.004	-0.009
		(0.005)	(0.006)	(0.006)	(0.007)
Ln(家庭人均住房价值)		0.016***	0.000	0.000	-0.007**
		(0.003)	(0.002)	(0.002)	(0.003)
少年抚养比		-0.740***	-0.397***	-0.409***	-0.302**
		(0.081)	(0.129)	(0.129)	(0.141)
老年抚养比		-0.315***	0.266**	0.257**	0.304***
		(0.082)	(0.106)	(0.106)	(0.107)
失业成员占比		0.273***	-0.017	-0.011	-0.017
		(0.050)	(0.051)	(0.051)	(0.056)
Ln(城市GDP)				0.169	0.171
				(0.175)	(0.173)
第三产业占比				0.329	0.076

				(0. 502)	(0. 578)
金融发展水平				0. 053***	0. 095***
				(0. 013)	(0. 026)
Ln(住房价格)				-0. 037	-0. 011
				(0. 033)	(0. 045)
Ln(数字普惠金融水平)				-0. 338	-0. 104
				(0. 277)	(0. 466)
Ln(家庭人均收入)的时间趋势					0. 009
					(0. 007)
Ln(家庭人均金融资产)的时间趋势					-0. 005
					(0. 004)
Ln(家庭人均住房价值)的时间趋势					-0. 005***
					(0. 002)
少年抚养比的时间趋势					0. 117**
					(0. 050)
老年抚养比的时间趋势					0. 055*
					(0. 029)
失业成员占比的时间趋势					-0. 015
					(0. 023)
是否拥有住房的时间趋势					-0. 057*
					(0. 030)
是否有住房负债的时间趋势					-0. 029
					(0. 021)
是否自营工商业的时间趋势					-0. 037*
					(0. 019)
受访者及配偶中养老保险人数为1的时间趋势					0. 055**
					(0. 027)
受访者及配偶中养老保险人数为2的时间趋势					-0. 013
					(0. 028)
受访者及配偶中医疗保险人数为1的时间趋势					-0. 136***
					(0. 042)
受访者及配偶中医疗保险人数为2的时间趋势					-0. 033
					(0. 038)
Ln(城市GDP)的时间趋势					0. 007
					(0. 010)
第三产业占比的时间趋势					-0. 134
					(0. 130)
金融发展水平的趋势					0. 025**
					(0. 012)
Ln(住房价格)的时间趋势					0. 005
					(0. 017)
常数项	10. 324***	7. 004***	5. 769***	4. 113	-79. 683
	(1. 664)	(0. 628)	(1. 195)	(4. 245)	(272. 323)
样本数	12804	12804	12804	12804	12804

注：本表为正文中表 2 的详细结果，为节约篇幅，只报告了连续变量的结果。以离散整数表示的变量均以虚拟变量的形式放入模型，部分结果未报告，包括：是否拥有住房、是否有住房负债、是否自营工商业、风险偏好、性别、年龄、受教育年限、婚姻状况、健康状况、户籍类型、城市虚拟变量、年份虚拟变量。时间趋势是指家庭及城市级别变量基期值的线性时间趋势。所有结果均使用家庭抽样权重。括号内为经过城市层面聚类调整的稳健标准误。*、**、*** 分别表示 10%、5% 和 1% 的显著性水平。

附录V 稳健性检验

表V1报告了一系列稳健性检验的结果。^①因为CHFS数据没有调查各种耐用品的购买年份，无法将使用支出分摊（李江一，2018）。表V1第（1）列中将家庭的消费支出减去耐用品支出再计算平均消费倾向，以此作为被解释变量来检验主要结果在回避上述分摊问题后的稳健性。结果较为稳健。

本文关注居民的实际消费支出，为排除价格变动的影响，如前文所述，我们使用省级居民消费价格指数将名义变量转换为实际变量。省内城市间的物价水平往往也具有差异，因此我们收集了样本期内各城市的居民消费价格指数。^②表V1第（2）列报告了使用城市居民消费价格指数对名义变量进行平减后的结果。相较于基准估计值，结果非常稳健。

为了考察结果在行政数据中的稳健性，在表V1第（3）列中，我们将基于LandScan计算的市辖区常住人口密度替换为基于地级市常住人口行政数据计算的人口密度。^③结果相较于基准估计值是稳健的。

考虑到我国城市人口主要分布于中心城区，特别是某些单中心的特大城市，本文的人口密度计算方法可能低估了这些城市的人口集聚程度。这可能会造成测量误差的问题，估计值将产生趋零偏误。正文中的工具变量结果表明，使用外生冲击产生的人口密度变化所估计的消费效应并未与OLS的估计结果有显著差异。这在一定程度上降低了测量误差可能造成显著偏误的疑虑。严谨起见，我们将一些人口密度可能被低估的特异性城市从样本中剔除，包括北京、上海、广州、深圳、重庆、天津。此外，这些城市在政策上可能也具有特异性。表V1第（4）列显示估计结果依然是稳健的，虽然在数值上有所增大，但不能判断其与主要估计结果间是否具有显著的差异。

为进一步降低测量误差，表V1第（5）列中剔除了没有与家庭主要成员居住在同一区县的样本。考虑到年龄过大的户主未必是家庭消费活动的主要决策者，表V1第（6）列剔除了户主年龄大于70岁的样本。人口规模过大的家庭特征、偏好以及消费行为可能与一般家庭不同，表V1第（7）列剔除了家庭人口数大于8的样本。出于对数据质量的考虑，表V1第（8）列仅使用户主或户主配偶受访的样本。上述估计结果均稳健。

表V1 稳健性检验（家庭平均消费倾向）

	(1)	(2)	(3)	(4)
	减去耐用品支出	城市价格指数	行政数据	剔除特异性城市
Ln(人口密度)	0.418*** (0.126)	0.443*** (0.135)	0.353** (0.150)	0.459*** (0.114)
样本数	10884	9971	10883	7736
	(5)	(6)	(7)	(8)
	剔除在外地居住样本	剔除户主年龄大于70 样本	剔除家庭人口数大于 8样本	剔除非户主或配偶受 访样本
Ln(人口密度)	0.475*** (0.147)	0.402*** (0.151)	0.478*** (0.133)	0.434*** (0.145)
样本数	9579	9800	10824	9975

^① 出于叙述方便的考虑，本文仅对表V1中以家庭平均消费倾向为被解释变量的结果进行说明。表V2中展示了家庭人均消费支出的结果。

^② 由于某些城市存在数据缺失，我们还是使用省级的居民消费价格指数进行基准分析。

^③ 我们也参考钟粤俊等（2020）的做法，使用城市市辖区的建成区就业密度（二、三产就业总人口/建成区面积）表征人口密度，结果依然证实人口集聚对家庭人均消费有显著正向影响。

注：被解释变量除第（1）列为家庭总支出减去耐用品支出后计算的家庭平均消费倾向以外，其余列均为原始的家庭平均消费倾向。表中结果均使用正文中表 2 第（4）列中的模型设定得出，具体的变量、样本设定请见正文说明。括号内为经过城市层面聚类调整的稳健标准误。**和***分别表示 5%和 1%的显著性水平。

表 V 2 稳健性检验（家庭人均消费支出的自然对数）

	(1)	(2)	(3)	(4)
	减去耐用品支出	城市价格指数	行政数据	剔除特异性城市
Ln(人口密度)	0.375*** (0.134)	0.377*** (0.133)	0.319** (0.138)	0.386*** (0.130)
样本数	12804	11706	12804	9196
	(5)	(6)	(7)	(8)
	剔除在外地居住样本	剔除户主年龄大于70 样本	剔除家庭人口数大于 8样本	剔除非户主或配偶受 访样本
Ln(人口密度)	0.402*** (0.128)	0.312** (0.152)	0.386*** (0.135)	0.347** (0.141)
样本数	11146	11519	12726	11726

注：被解释变量除第（1）列为家庭总支出减去耐用品支出后计算的家庭人均消费支出的自然对数以外，其余列均为原始的家庭人均消费支出的自然对数。表中结果均使用正文中表 2 第（4）列中的模型设定得出，具体的变量、样本设定请见正文说明。括号内为经过城市层面聚类调整的稳健标准误。**和***分别表示 5%和 1%的显著性水平。

附录VI 异质性分析

表VI1 报告了异质性分析结果。^①首先,我们考察人口集聚是否对居民消费结构有影响。我们将生存型消费、发展型消费和享受型消费分别占家庭消费支出的比重作为被解释变量。^②表VI1 的A 栏第(1) - (3) 列的估计结果表明,人口集聚降低了居民生存型消费占比和发展型消费占比,但在统计意义上不显著;而人口密度提高 1%,享受型消费占比的值在 5%的统计水平上显著提高 0.072 个百分点。伴随着城市人口密度的提升,居民对需求弹性较大的商品以及文化娱乐、旅游等服务消费的需求提升,生活质量提高。这表明人口集聚不仅增加了居民消费支出,而且促进了消费升级。政策方面,应加强供给侧改革,促进服务业发展,以适应居民消费结构的变化。

接下来,我们进一步考察人口集聚对不同收入群体消费支出的影响,表VI1 的B 栏报告了回归结果。^③结果表明,人口集聚对不同收入的家庭在平均消费倾向上具有异质性的影响。对于低收入家庭,人口集聚对消费的促进效应在 5%的水平上显著。人口集聚对中等收入家庭平均消费倾向的影响为正,但并不显著。对于高收入家庭,人口集聚对家庭平均消费倾向具有显著的促进效应。人口集聚程度提高 1%,低收入家庭和高收入家庭的平均消费倾向分别增加 0.957 和 0.578 个百分点。表VI2 中展示的以家庭人均消费支出自然对数作为被解释变量的对应结果显示,人口集聚对消费支出的促进作用主要集中于中高收入家庭。该结果与上述人口集聚促进享受型消费的结果是相关的,虽然人口集聚会提高低收入家庭的平均消费倾向,但从支出上看,中高收入家庭进行消费升级的倾向更大。因此,一方面,以人口集聚促进消费的思路制定相关政策时,应充分考虑政策目标群体间的差异性。此外,还应充分考虑排序效应对政策效果的可能影响,尤其是什么因素导致了高(低)消费人群在低(高)密度地区的非随机分布。这是政策制定者需要知道的,因此在后续研究中应进行充分考察。另一方面,政策制定者需要意识到相关政策可能导致的消费不平等。因此,后续研究还应关注促进低收入家庭在高密度地区消费的因素。

最后,我们考虑人口集聚的消费效应在城市规模上的异质性,表VI1 的C 栏报告了分析结果。^④我们发现,人口集聚的消费效应主要体现在大城市样本中。人口集聚对中小城市的居民消费倾向有正向影响但在统计意义上不显著。在大城市中,人口密度提高 1%,家庭平均消费倾向在 5%的统计水平上增加 0.335 个百分点。对于特大或超大城市,估计值为正但不显著。这意味着,人口集聚的消费效应在城市规模上呈现出先增加后降低的非线性。因此,在利用人口集聚的消费效应进行政策制定时还应考虑其不同城市规模上的有效性差异。对于特大或超大城市,政策应倾向于疏解人口过多可能造成的负外部性;对于大城市,应充分利用显著的人口集聚的消费效应带动经济的良性循环发展;对于中小城市而言,政策的首要目标可能是以增加城市规模为先导。此外,上述城市规模异质性意味着可以制定不同规模城市间的人口迁移政策来实现整体消费净增长的可能性。例如,鼓励超大城市人口向中小城市转移,一方面疏解了超大城市的承载压力,缓解负外部性对消费的影响,另一方面中小城市人口规模的增加也会使人口集聚的消费效应提升,从而实现整体消费的净增长。然而,从超

^① 出于叙述方便的考虑,本文仅对表VI1 中以家庭平均消费倾向为被解释变量的结果进行说明。家庭人均消费支出的结果见表VI2。

^② 生存型消费包括:食品支出、衣着支出、日用品支出、居住支出;发展型消费包括:交通支出、通信支出、教育支出、医疗保健支出;享受型消费包括:家政支出、耐用品支出、文化娱乐支出、旅游支出、奢侈品支出。

^③ 本文以 CHFS2017 年的收入数据为标准,将年收入最低 30%的家庭划为低收入家庭;收入最高的 30%家庭划为高收入家庭;其余为中等收入家庭。

^④ 本文以 2015 年城市市辖区的人口规模为标准,将 100 万人以下的城市定义为中小城市,将 100 至 500 万人的城市定义为大城市,将 500 万人以上的城市定义为特大或超大城市。

大城市迁往小城市的人口是否到小城市后会改变消费决策还缺乏经验证据。集聚的消费效应对迁入人口来说是水平效应还是增长效应这一问题需要后续研究提供更多的实证证据。

表 VI1 异质性分析（家庭平均消费倾向）

A. 消费结构	被解释变量		
	生存型消费占比	发展型消费占比	享受型消费占比
	(1)	(2)	(3)
Ln(人口密度)	-0.030 (0.042)	-0.090 (0.082)	0.072** (0.032)
样本数	12804	12804	12804
被解释变量：家庭平均消费倾向			
B. 家庭收入	低收入家庭	中等收入家庭	高收入家庭
	(1)	(2)	(3)
Ln(人口密度)	0.957** (0.426)	0.218 (0.189)	0.578*** (0.168)
样本数	1859	4921	4103
被解释变量：家庭平均消费倾向			
C. 城市规模	中小城市	大城市	特大或超大城市
	(1)	(2)	(3)
Ln(人口密度)	0.117 (0.236)	0.335** (0.131)	0.226 (0.230)
样本数	1657	4517	4709

注：A 栏的被解释变量为家庭生存型、发展型、享受型消费占家庭消费的比重。B、C 栏的被解释变量为家庭平均消费倾向。A 栏每列使用所有样本。B 栏分别使用低、中、高收入家庭样本。C 栏分别使用中小城市、大城市以及特大或超大城市的家庭样本。除了上述被解释变量或样本不同，模型设定均与正文中表 2 第（4）列一致。结果均使用家庭抽样权重。括号内为经过城市层面聚类调整的稳健标准误。**和***分别表示 5%和 1%的显著性水平。

表 VI2 异质性分析（家庭人均消费支出的自然对数）

A. 家庭收入	被解释变量：Ln(家庭人均消费)		
	低收入家庭	中等收入家庭	高收入家庭
	(1)	(2)	(3)
Ln(人口密度)	-0.053 (0.333)	0.397** (0.183)	0.479*** (0.152)
样本数	2528	5394	4882
被解释变量：Ln(家庭人均消费)			
B. 城市规模	中小城市	大城市	特大或超大城市
	(1)	(2)	(3)
Ln(人口密度)	-0.451 (0.431)	0.360** (0.136)	0.351** (0.125)
样本数	2006	5394	5404

注：A、B 栏的被解释变量为家庭人均消费的自然对数。A 栏分别使用低、中、高收入家庭样本。B 栏分别使用中小城市、大城市、特大或超大城市的家庭样本。除上述被解释变量或样本不同，模型设定均与正文中表 2 第（4）列一致。结果均使用家庭抽样权重。括号内为经过城市层面聚类调整的稳健标准误。**和***分别表示 5%和 1%的显著性水平。

附录VII 人口集聚与不可贸易服务密度

如附录 I 中所述，商品特征距离的期望 $E \sum_{Y(x)} |x - y| = \theta^2 / 4n$ 随人口密度的增加而减小是基于 θ 不变的假设。在此假设下易知 $\frac{dm}{m} / \frac{dn}{n} = 1$ ，即不可贸易服务密度与人口密度的弹性为单位弹性，不存在规模效应。表 VII1 中报告的结果验证了这一假设。若使用不可贸易服务密度作为消费匹配机会的代理变量，这一结果也与理论框架所述相符，即大部分文献得到的总体匹配函数是规模报酬不变的。从结果来看，人口密度提升了消费匹配的机会，虽然没有体现出规模递增的效应，但总的来说，人口密度的增加在一定程度上降低了消费的交易成本，从而促进居民的消费。

表 VII1 人口集聚对不可贸易服务密度的影响

	被解释变量					
	Ln(不可贸易服	Ln(餐饮服务密	Ln(购物服务密	Ln(科教文化服	Ln(生活服务密	Ln(体育休闲服
	务密度)	度)	度)	务密度)	度)	务密度)
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(人口密度)	1.010** (0.431)	1.244* (0.661)	1.141** (0.563)	1.057*** (0.196)	0.752** (0.330)	1.219*** (0.206)
单位弹性检验P值	0.982	0.713	0.803	0.773	0.455	0.290
样本数	166	166	166	166	166	166

注：各类不可贸易服务密度是城市市辖区各类兴趣点数量总和与市辖区建成区面积之比。控制变量包括：GDP 规模、第三产业占 GDP 比重、金融发展水平、住房价格、数字普惠金融、城市虚拟变量、年份虚拟变量。括号内为经过城市层面聚类调整的稳健标准误。单位弹性检验的原假设为：Ln(人口密度)的系数估计值等于 1。*、**、***分别表示 10%、5%、1%的显著性水平。

附录VIII 住房支出估算

CHFS2017 会询问部分家庭“目前，这套房子如果出租，一个月能收多少租金？”以及“目前，这所房子值多少钱？”，利用受访者回答的租金除以售价，得到单套住房的租售比。由于部分家庭存在多套房的情况（CHFS2017 至多会询问受访者六套房的信息），同样得到多套住房的租售比。在此基础上，利用调查权重计算得到在城市层面的加权平均租售比。存在缺失的城市使用其余城市的均值来代替。

考虑到城市的住房市场特征在短期内不会发生太大变化，假定城市平均租售比在 2013 至 2017 年保持一致。然后，将城市租售比与家庭自住住房的现值相乘，近似得到住房自住家庭的住房支出。

参考文献

- [1] Coles, M.G., “Understanding the Matching Function: The Role of Newspapers and Job Agencies”, *CEPR Discussion Papers*, 1994.
- [2] Coles, M.G., and E. Smith, “Marketplaces and Matching”, *International Economic Review*, 1998, 39(1), 239.
- [3] Duranton G., and D. Puga, “Chapter 48 - Micro-Foundations of Urban Agglomeration Economies”. Henderson J V, Thisse J-F, eds.//, *Handbook of Regional and Urban Economics*, Elsevier, 2004.
- [4] Gaigné, C., S. Riou, and J.-F. Thisse, “Are Compact Cities Environmentally Friendly?”, *Journal of Urban Economics*, 2012, 72(2), 123 - 136.
- [5] Glaeser, E.L., and J.D. Gottlieb, “The Economics of Place-Making Policies”, *Brookings Papers on Economic Activity*, 2008, 155 - 239.
- [6] Glaeser, E.L., and B. Sacerdote, “Why is There More Crime in Cities?”, *Journal of Political Economy*, 1999, 107(S6), S225 - S258.
- [7] 郭峰、王靖一、王芳、孔涛、张勋、程志云, “测度中国数字普惠金融发展: 指数编制与空间特征”, 《经济学(季刊)》, 2020 年第 4 期, 第 1401-1418 页。
- [8] Helsley, R.W., and W.C. Strange, “Matching and Agglomeration Economies in a System of Cities”, *Regional Science and Urban Economics*, 1990, 20(2), 189 - 212.
- [9] Kang, H., H. Suh, and J. Yu, “Does Air Pollution Affect Consumption Behavior? Evidence from Korean Retail Sales”, *Asian Economic Journal*, 2019, 33(3), 235 - 251.
- [10] 李江一, ““房奴效应”导致居民消费低迷了吗?”, 《经济学(季刊)》, 2018 年第 1 期, 第 405-430 页。
- [11] Li, Y., and X. Liu, “How did Urban Polycentricity and Dispersion Affect Economic Productivity? A Case Study of 306 Chinese Cities”, *Landscape and Urban Planning*, 2018, 173, 51—59.
- [12] Mejía, D., and P. Restrepo, “Crime and Conspicuous Consumption”, *Journal of Public Economics*, 2016, 135, 1 - 14.
- [13] 尹志超、吴雨、甘犁, “金融可得性、金融市场参与和家庭资产选择”, 《经济研究》, 2015 年第 3 期, 第 87-99 页。
- [14] 钟粤俊、陆铭、奚锡灿, “集聚与服务业发展——基于人口空间分布的视角”, 《管理世界》, 2020 年第 11 期, 第 35-49 页。

注：该附录是期刊所发表论文的组成部分，同样视为作者公开发表的内容。如研究中使用该附录中的内容，请务必在研究成果上注明附录下载出处。